Ministry of Higher Education and
Scientific Research
University of Mosul
College of Computer Science and Mathematics
Department of Computer Science

Utilizing Deep Learning Techniques For Electronic Health Records

**A Thesis Submitted to the Council of the College of
Computer Science and Mathematics
University of Mosul
as a Partial Fulfillment of Requirements for the Degree of Master
in
Computer Science**

By

Dalal Nawfal Hamid Ali

Supervised by
Assist. Prof. Dr. Mohammed Chachan Younis

2024 A.D. **1446 A.H.**

ABSTRACT

Given the great development in artificial intelligence algorithms, whose applications have entered many areas of life, facilitating the solution of many simple and complex issues, one of these applications is the innovation of a new intelligent technology based on deep learning to predict the duration of stay of diabetic patients in the hospital, after analyzing the data of patients with this disease according to the available electronic health records.

Extensive experiments have been conducted to compare different deep learning architectures, including fully connected neural networks (FCNNs), long short-term memory networks (LSTMs), simple recurrent neural networks (RNNs), and deep belief networks (DBNs). These models were trained on a large-scale dataset from EHRS. This study discusses the construction of a comprehensive workflow model for applying different AI tools to a clinical setting such as patient selection, drug testing, and predicting the expected days of hospital stay for a patient. FCNNs, LSTMs, RNNs, and DBNs are used in the diabetes electronic health record (EHR). The Root Mean Squared Propagation (RMSProp) algorithm is used in the training model as an optimizer to improve the training process by dynamically adjusting the learning rate and accelerating the training.

Experimental results demonstrated the superior performance of the fully connected neural network (FCNN) model compared to other architectures such as LSTM, simple RNN, and DBNs across different classification metrics. The FCNN model achieved the best performance across all compared metrics, including the highest accuracy value of about 0.735. In contrast, the other algorithms achieved lower accuracy values, with LSTM achieving 0.697, SimpleRNN achieving 0.601, and DBNs achieving 0.585. The study focuses on analyzing and testing data on

drugs used to treat diabetes patients (various tests) using AI techniques in healthcare to predict the length of hospital stay for diabetic patients, which contributes to improving the quality of healthcare and reducing costs.

وزارة التعليم العالي والبحث العلمي
جامعة الموصل
كلية علوم الحاسوب والرياضيات
قسم علوم الحاسوب

استخدام تقنيات التعلم العميق لسجلات الصحة الالكترونية

رسالة مقدمة
الى مجلس كلية علوم الحاسوب والرياضيات في جامعة الموصل
كجزء من متطلبات نيل شهادة ماجستير علوم في
علوم الحاسوب

من قبل

دلال نوفل حامد علي

بإشراف

أ.م.د. محمد جاجان يونس

١٤٤٦ هـ ٢٠٢٤ م

الخلاصة

في ظل التطور الكبير في خوارزميات الذكاء الاصطناعي، والتي دخلت تطبيقاتها العديد من مجالات الحياة، مما سهل حل العديد من القضايا البسيطة والمعقدة، أحد هذه التطبيقات هو ابتكار تقنية ذكائية جديدة تعتمد على التعلم العميق للتنبؤ بمدة إقامة مرضى السكري في المستشفى، بعد تحليل بيانات المرضى بهذا المرض وفقاً للسجلات الصحية الإلكترونية المتوفرة.

تم اجراء تجارب مكثفة لمقارنة بنيات التعلم العميق المختلفة، بما في ذلك الشبكات العصبية المتصلة بالكامل (FCNNs)، وشبكات الذاكرة القصيرة المدى الطويلة (LSTMs)، والشبكات العصبية المتكررة البسيطة (RNNs)، وشبكات المعتقدات العميقة (DBNs). تم تدريب هذه النماذج على مجموعة بيانات واسعة النطاق من EHRS. تناقش هذه الدراسة بناء نموذج سير عمل شامل لتطبيق أدوات الذكاء الاصطناعي المختلفة على بيئة سريرية مثل اختيار المريض، واختبار الأدوية، والتنبؤ بالأيام المتوقعة لبقاء المريض في المستشفى. استخدام FCNNs وLSTMs وRNNs وDBNs في السجل الصحي الإلكتروني لمرض السكري (EHR). وتم استخدام خوارزمية RMSProp في نموذج التدريب كمحسن إلى تحسين عملية التدريب عن طريق ضبط معدل التعلم ديناميكيًا وتسريع التدريب.

وأظهرت النتائج التجريبية الأداء المتفوق لنموذج الشبكة العصبية المتصلة بالكامل (FCNN) مقارنة بالهياكل الأخرى مثل LSTM و Simple RNN و DBNs عبر مقاييس التصنيف المختلفة. حقق نموذج FCNN أفضل أداء عبر جميع المقاييس المقارنة، بما في ذلك أعلى قيمة دقة تبلغ حوالي 0.735. في المقابل، حققت الخوارزميات الأخرى قيم دقة أقل، حيث حقق LSTM 0.697 و SimpleRNN 0.601 و DBNs 0.585. تركز الدراسة على تحليل واختبار البيانات حول الأدوية المستخدمة لعلاج مرضى السكري (اختبارات مختلفة) باستخدام تقنيات الذكاء الاصطناعي في الرعاية الصحية للتنبؤ بطول مدة إقامة مرضى السكري في المستشفى، مما يساهم في تحسين جودة الرعاية الصحية وخفض التكاليف.